

T S4/5/1

4/5/1

DIALOG(R)File 351:Derwent WPI

(c) 2004 Thomson Derwent. All rts. reserv.

010331867 **Image available**

WPI Acc No: 1995-233559/199531

XRPX Acc No: N95-182052

Speech recognition system for telephone network - determines identity of caller and processes signal against caller specific library

Patent Assignee: INT BUSINESS MACHINES CORP (IBMC); IBM CORP (IBMC)

Inventor: COHEN P S; LUCASSEN J M; SHERWIN E B; VIZCAINO J L

Number of Countries: 005 Number of Patents: 003

Patent Family:

Patent No	Kind	Date	Applicat No	Kind	Date	Week
EP 661690	A1	19950705	EP 94309591	A	19941220	199531 B
JP 7210190	A	19950811	JP 94304072	A	19941207	199541
US 5802251	A	19980901	US 93175707	A	19931230	199842
			US 95523755	A	19950905	

Priority Applications (No Type Date): US 93175707 A 19931230; US 95523755 A 19950905

Cited Patents: EP 302614; EP 549265; US 5033089

Patent Details:

Patent No	Kind	Lan	Pg	Main IPC	Filing Notes
-----------	------	-----	----	----------	--------------

EP 661690	A1	E	9	G10L-009/00	
-----------	----	---	---	-------------	--

Designated States (Regional): DE FR GB

JP 7210190	A		9	G10L-003/00	
------------	---	--	---	-------------	--

US 5802251	A			G10L-003/00	Cont of application US 93175707
------------	---	--	--	-------------	---------------------------------

Abstract (Basic): EP 661690 A

The speech recognition system includes a store for a core library of speech templates. Several caller-specific libraries of speech templates are stored. An identification of a caller within a telephone network is determined. An input speech utterance is processed against the core library in an event that identification cannot be determined.

An input speech utterance is processed against a selected caller specific libraries when an identity of the caller is determined. The caller is identified using an identification system within the network. The caller specific libraries are processed to reflect variations in speech utterance.

ADVANTAGE - Improved sound recognition accuracy due to taking into account regional vocabulary and accents.

Dwg.1/3

Title Terms: SPEECH; RECOGNISE; SYSTEM; TELEPHONE; NETWORK; DETERMINE; IDENTIFY; CALL; PROCESS; SIGNAL; CALL; SPECIFIC; LIBRARY

Derwent Class: P86; W01; W04

International Patent Class (Main): G10L-003/00; G10L-009/00

International Patent Class (Additional): G10L-005/06; G10L-007/08;

G10L-009/06; H04Q-007/38

File Segment: EPI; EngPI

THIS PAGE BLANK (USPTO)

JPA07-210190 which corresponds to USP 5,802,251

(19) 日本国特許庁 (J P)

(12) 公開特許公報 (A)

(11) 特許出願公開番号

特開平7-210190

(43) 公開日 平成7年(1995)8月11日

(51) Int. Cl. °	識別記号	庁内整理番号	F I	技術表示箇所
G10L 3/00	551 A			
	531 K			
H04Q 7/38		7605-5K	H04B 7/26	109 S

審査請求 有 請求項の数 8 OL (全9頁)

(21) 出願番号 特願平6-304072

(22) 出願日 平成6年(1994)12月7日

(31) 優先権主張番号 175707

(32) 優先日 1993年12月30日

(33) 優先権主張国 米国 (US)

(71) 出願人 390009531

インターナショナル・ビジネス・マシーンズ・コーポレーション

INTERNATIONAL BUSINESS MACHINES CORPORATION

アメリカ合衆国10504、ニューヨーク州アーモンク (番地なし)

(72) 発明者 ポール・エス・コーアン

アメリカ合衆国10598 ニューヨーク州ヨークタウン・ハイツ ナットリー・サークル 3271

(74) 代理人 弁理士 合田 潔 (外2名)

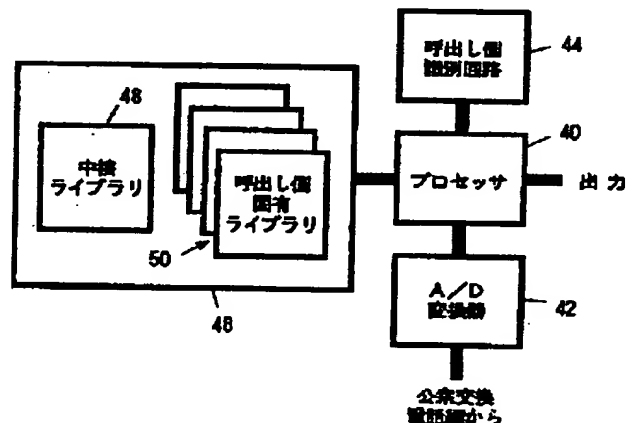
最終頁に続く

(54) 【発明の名称】 音声認識方法及びシステム

(57) 【要約】 (修正有)

【目的】 判別された呼出し側のIDに基づいて電話網内の音声認識システムの混乱を削減するための方法およびシステム。

【構成】 音声認識システムにおいて、音声の基本語彙を表す音声テンプレート48、複数の呼出し側固有音声テンプレート50 (具体的な地理的位置と特定の個人に対応する専用語彙と発音を表す音声テンプレートを含む) が作成され、格納される。電話網を介してシステムに持ち込まれる帯域幅の縮小、伝送チャネルの変動、およびその他の信号変化を反映する処理をすることが好ましい。呼出し側のIDは、ネットワークへの接続時に標準の呼出し側識別回路44を介して判別され、発声を検出すると、呼出し側のIDを判別できない場合は中核ライブラリ48と照らし合わせてその発声进行处理し、呼出し側のIDを判別できる場合は特定の呼出し側固有ライブラリ50と照らし合わせてその発声进行处理する。



【特許請求の範囲】

【請求項1】電話網を介して音声を表す格納済みテンプレートと照らし合わせて音声の入力フレームを処理する音声認識システムの精度と効率を高めるための方法において、

音声テンプレートの中核ライブラリを作成して格納するステップと、

複数の呼出し側固有音声テンプレート・ライブラリを作成して格納するステップと、

前記電話網内の呼出し側のIDを判別しようと試みるステップと、

前記電話網内の前記呼出し側のIDが判別されない場合に、音声テンプレートの前記中核ライブラリと照らし合わせて入力音声発声処理するステップと、

前記電話網内の前記呼出し側のIDの判別に対する応答として、前記複数の呼出し側固有音声テンプレート・ライブラリから選択されたライブラリと照らし合わせて入力音声発声処理するステップとを含む方法。

【請求項2】前記電話網内の呼出し側のIDを判別する前記ステップが、前記電話網内の呼出し側識別システムを使用して前記呼出し側のIDを判別するステップを含むことを特徴とする、請求項1に記載の電話網を介して音声を表す格納済みテンプレートと照らし合わせて音声の入力フレームを処理する音声認識システムの精度と効率を高めるための方法。

【請求項3】複数の呼出し側固有音声テンプレート・ライブラリを作成して格納する前記ステップが、前記電話網内での伝送の結果として発生する音声発声の変化を反映するように処理される複数の呼出し側固有音声テンプレート・ライブラリを作成して格納するステップを含むことを特徴とする、請求項1に記載の電話網を介して音声を表す格納済みテンプレートと照らし合わせて音声の入力フレームを処理する音声認識システムの精度と効率を高めるための方法。

【請求項4】複数の呼出し側固有音声テンプレート・ライブラリを作成して格納する前記ステップが、具体的な地理的位置を反映する語彙および発音を含む複数の呼出し側固有音声テンプレート・ライブラリを作成して格納するステップを含むことを特徴とする、請求項1に記載の電話網を介して音声を表す格納済みテンプレートと照らし合わせて音声の入力フレームを処理する音声認識システムの精度と効率を高めるための方法。

【請求項5】電話網を介して音声を表す格納済みテンプレートと照らし合わせて音声の入力フレームを処理する音声認識システムの精度と効率を高めるためのシステムにおいて、

音声テンプレートの中核ライブラリを作成して格納するための手段と、

複数の呼出し側固有音声テンプレート・ライブラリを作成して格納するための手段と、

前記電話網内の呼出し側のIDを判別しようと試みるための手段と、

前記電話網内の前記呼出し側のIDが判別されない場合に、音声テンプレートの前記中核ライブラリと照らし合わせて入力音声発声処理するための手段と、

前記電話網内の前記呼出し側のIDの判別に対する応答として、前記複数の呼出し側固有音声テンプレート・ライブラリから選択されたライブラリと照らし合わせて入力音声発声処理するための手段とを含むシステム。

【請求項6】前記電話網内の呼出し側のIDを判別するための前記手段が、前記電話網内の呼出し側識別システムを使用して前記呼出し側のIDを判別するための手段を含むことを特徴とする、請求項5に記載の電話網を介して音声を表す格納済みテンプレートと照らし合わせて音声の入力フレームを処理する音声認識システムの精度と効率を高めるためのシステム。

【請求項7】複数の呼出し側固有音声テンプレート・ライブラリを作成して格納するための前記手段が、前記電話網内での伝送の結果として発生する音声発声の変化を反映するように処理される複数の呼出し側固有音声テンプレート・ライブラリを作成して格納するための手段を含むことを特徴とする、請求項5に記載の電話網を介して音声を表す格納済みテンプレートと照らし合わせて音声の入力フレームを処理する音声認識システムの精度と効率を高めるためのシステム。

【請求項8】複数の呼出し側固有音声テンプレート・ライブラリを作成して格納するための前記手段が、具体的な地理的位置を反映する語彙発音を含む複数の呼出し側固有音声テンプレート・ライブラリを作成して格納するための手段を含むことを特徴とする、請求項5に記載の電話網を介して音声を表す格納済みテンプレートと照らし合わせて音声の入力フレームを処理する音声認識システムの精度と効率を高めるためのシステム。

【発明の詳細な説明】

【0001】

【産業上の利用分野】本発明は、一般に、改良された音声認識システムに関し、特に、音声認識精度を高めるための改良された方法およびシステムに関する。より具体的には、本発明は、呼出し側固有音声テンプレート・ライブラリを使用して電話網内の音声認識を高めるための方法およびシステムに関する。

【0002】

【従来の技術】音声認識は先行技術で周知の技術である。既知の話し手用の所与の語彙から分離した単語を認識することが、おそらく最も単純なタイプの音声認識であり、この種の音声認識がしばらく前から知られていた。一般に、認識対象語彙内の単語は個々のテンプレートとして事前記憶され、それぞれのテンプレートが語彙内の単語の音声パターンを表している。分離した単語が話されると、システムは、この単語と、語彙を表す個々

のテンプレートそれぞれとを比較するだけである。この方法は一般に全単語テンプレート突合せ（マッチング）と呼ばれている。成功した多くの音声認識システムでは、この方法とともに動的プログラミングを使用して、話された単語と事前記憶テンプレートとの非線形のタイム・スケールの変化に対処している。

【0003】連続音声や、固有名詞または地名を含む音声認識するのはもっと難しい。連続音声、すなわち、単語の連続は、先行技術では複数経路動的プログラミングを使用して認識されていた。このようなシステムの一
10 例は、H. Sakoeによる"Two Level DP Matching A Dynamic Programming Based Pattern Matching Algorithm For Connected Word Recognition" (IEEE Transactions on Acoustics Speech and Signal Processing, Volume AS-SP-27, No. 6, pp. 588-595, 1979年12月) に提案されている。この論文では、入力パターン全体に最も一致する一連の単語テンプレートを見つけるための2パス動的プログラミング・アルゴリズムを提案している。システムを通るそれぞれのパスは、入力パターンのすべての可能な部分と突き合わせたすべてのテンプレートに
20 における類似性を示すスコアを生成する。もう1つのパスでは、このスコアを使用して、入力パターン全体に対応する最良のテンプレート・シーケンスを見つける。

【0004】米国特許第5040127号では、入力フレームと、音声を表す事前記憶テンプレートとを比較し、認識される可能性のある個々の単語として検討中の各テンプレートごとにリンク・ネットワーク内のレコード間にリンクを確立することによって連続音声を処理する、連続音声認識システムを提案している。リンクされたレコードは、索引付きデータ・セットとして格納され
30 ている親リンク・レコードと子孫リンク・レコードを含み、各データ・セットは、テンプレートを表す記号と、リンク・レコードが格納された相対時間を表す手順標識と、その発生元であるネットワーク内のリンク・レコードを示すポインタとを含む。

【0005】固有名詞の認識は、音声認識システムにとっていわゆる「混乱」が増加することを意味し、この問題は最近、米国特許第5212730号で認識されている。この特許は、複数の発音が可能な固有名詞の発声表現を認識するためのテキスト派生認識モデルを使用して
40 名前の認識を行う。この特許に記載されている名前認識方法は、名前テキストを指定してから、名前テキストから指定の数のテキスト派生認識モデルを構築することでアクセスされるテキスト・データベースへの名前テキストの入力を含む。この場合、各テキスト派生認識モデルは、その名前の少なくとも1つの発音を表す。その後、発声した名前入力によってテキスト・データベースへのアクセスを試みるたびに、テキスト・データベースと発声した名前入力とを比較し、一致が可能かどうかを判別する。

【0006】米国特許第5202952号では、発声をフレーム・データ・セットに変換することで音声を認識する大規模語彙連続音声事前選別処理システムを開示している。この場合、各フレーム・データ・セットが平滑されて、所定の数のフレームにおける平滑フレーム・モデルを生成する。一続きのフレーム期間において音響的に類似の単語モデルのクラスタを常駐語彙として指定し、平滑フレーム・モデルとの比較対象となるクラスタ用の確率モデルを使用して評価した平滑フレームの可能性を含むクラスタ・スコアをシステムが生成する。

【0007】これらのシステムはどれも、音声認識を正しく行うには連続音声発声の混乱を削減する必要があると認識している。この問題を取り扱っている文献としては、"Perplexity-A Measure of Difficulty of Speech Recognition Tasks" (Journal of the Acoustical Society of America, Volume 62, Supplement No. 1, p. S-63, 1977年秋号) と、"Continuous Speech Recognition Statistical Methods" (Handbook of Statistics Volume 2: Classification, Pattern Recognition and Reduction of Dimensionality, pp. 549-573, North-Holland Publishing Company, 1982年) がある。

【0008】上記の説明を考慮し、音声認識を正しく行うには、固有名詞、地名、および数値の場合に特に困難な問題である、発音が似ている多数の単語間を区別する能力を高める必要があることに留意されたい。したがって、音声認識の精度と効率を高める方法およびシステムが要求されていることに留意されたい。

【0009】

【発明が解決しようとする課題】本発明の一目的は、改良された音声認識システムを提供することにある。

【0010】本発明の他の目的は、音声認識の精度と効率を高めるための改良された方法およびシステムを提供することにある。

【0011】本発明の他の目的は、呼出し側固有音声テンプレート・ライブラリを使用して電話網内の音声認識を高めるための改良された方法およびシステムを提供することにある。

【0012】

【課題を解決するための手段】上記の諸目的は、以下に説明する通りに達成される。判別された呼出し側のIDに基づいて電話網内の音声認識システムの混乱を削減するための方法およびシステムを開示する。音声を表す格納済みテンプレートと照らし合わせて音声の入力フレームを処理する音声認識システムでは、音声の基本語彙を表す音声テンプレートの中核ライブラリが作成され、格納される。複数の呼出し側固有音声テンプレート・ライブラリも作成され、格納されるが、それぞれのライブラリは、具体的な地理的位置と特定の個人に対応する専用語彙を表す音声テンプレートを含む。さらに、電話網を介してシステムに持ち込まれる帯域幅の縮小、伝送チャ
50

ネルの変動、およびその他の信号変化を反映するために、呼出し側固有音声テンプレート・ライブラリを処理することが好ましい。呼出し側のIDは、ネットワークへの接続時に標準の呼出し側識別回路を介して判別され、発声を検出すると、呼出し側のIDを判別できない場合は中核ライブラリと照らし合わせてその発声を処理し、呼出し側のIDを判別できる場合は特定の呼出し側固有ライブラリと照らし合わせてその発声を処理し、それにより、システムによる音声認識の効率と精度が大幅に向上する。

【0013】

【実施例】ここで、添付図面、特に、図1を参照して説明すると、同図には、本発明の方法およびシステムを実現するために使用できる分散電話網が示されている。図示の通り、公衆交換電話網10を介して複数のユーザ位置がホスト位置12に連結されている。公衆交換電話網10は、このような通信を実現するための周知の技術のいずれかを使用して、電話通信を介して複数のユーザをホスト位置12に連結するよう機能することが好ましい。たとえば、ユーザ位置14は、通信チャンネル32、公衆交換電話網10、および通信チャンネル30を介してホスト位置12に連結された標準的な電話機18の使用を反映している。電話機18を使用してユーザが入力した音声は、コンピュータ16を使用して実現した音声認識システムを使って認識することができる。コンピュータ16は、IBM PS/2コンピュータなどのいわゆる「パーソナル」コンピュータのような適当なコンピュータを使用して実現できる。

【0014】あるいは、図1に示すように、ユーザは、無線塔22による無線周波伝送を介して通信するセルラ方式の移動電話20を使用することもできる。通常、無線塔22は、陸上通信チャンネル34を使用して公衆交換電話網10に連結される。さらに、最新の大陸横断通信は、衛星26および衛星受信機24を使って図示されているような衛星通信を使用して実現される場合も多い。この場合、衛星受信機24は、通信チャンネル36を介して公衆交換電話網10に連結される。

【0015】図1に示すように、最新の分散電話網は、ユーザがホスト位置12との通信を確立できるようにするための複数の多様な通信チャンネルを提供する。このような通信チャンネルは、コンピュータ16を使用して実現される音声認識システムの精度に影響する要素がそれぞれ明らかに異なっている。たとえば、所与の通信チャンネルでは帯域幅が縮小されている可能性がある。衛星システムは、伝送エコーや信号遮断の諸問題に悩まされる場合もある。しかも、信号品質が予測不能であること、様々な電話機のマイク特性が不明であること、および様々な地域アクセントも、図1に示すような分散電話網を使用した音声認識システムを実現する上での困難の一因となる。

【0016】また、公衆交換電話網内で選択した通信チャンネルは、発声の質および内容を変化させる既知の圧縮アルゴリズムやその他の各種信号処理技術を使用している場合が多く、その発声の認識がローカル・システム内での音声認識より困難になっている。

【0017】次に図2を参照すると、同図には、図1のコンピュータ16を使用して実現できる音声認識システムの高水準ブロック図が示されている。このシステムは、呼出し側のIDを使用して、このようなシステムの音声認識の混乱を削減する方法を示している。図2に示すように、音声の基本語彙を表す音声テンプレートの中核ライブラリ48を含むメモリ46が、コンピュータ16内に実現された音声認識システム内部に設けられている。同様に、メモリ46内には複数の呼出し側固有ライブラリ50も格納されている。それぞれの呼出し側固有ライブラリ50は、その呼出し側が通常使用する通信チャンネルに関連する、特定の地理的位置に関連する専用語彙を表すテンプレートを含むことが好ましく、これらのテンプレート内のデータは、帯域幅、マイク特性、アナログ信号品質、および図1の分散電話網内の特定の呼出し側に関連するその他の各種パラメータを反映するよう変更されていることが好ましい。

【0018】当業者は、このような呼出し側固有ライブラリが、発声の検出に必要な通信チャンネルをモデリングするネットワークを介して発声を選別し処理することで作成できることに留意されたい。しかも、本明細書を参照し、それぞれの呼出し側固有ライブラリが、分散電話網内の指定の呼出し側のIDに関連する指定の地理的位置に密接に関係する具体的な地理的位置、事業所、または固有名詞を表す一連の音声テンプレートを含んでもよいことに留意されたい。

【0019】したがって、ユーザと、コンピュータ16内に実現された音声認識システムとの間で通信が確立し、音声発声が検出されるたびに、その発声がアナログ／デジタル変換器42を使用して処理用に適切に変換され、プロセッサ40に連結されることが好ましい。次に、プロセッサ40は、呼出し側識別回路44と関連して公衆交換電話網から得られる呼出し側識別信号を使用して、使用電話機を識別することで特定の呼出し側のIDを確立する。当業者が理解できるように、このIDにより、呼出し側とホスト位置との間で検出される典型的な通信経路に基づいて、特定の呼出し側の地理的位置とその特定の呼出し側に通常関連する通信チャンネル・パラメータに関する有用な情報が得られる。

【0020】このため、呼出し側識別回路44の出力は、メモリ46内に含まれる複数の呼出し側固有ライブラリ50のうちの特定のライブラリをプロセッサ40が選択できるようにするためにプロセッサ40が使用する。次に、音声データの入力フレームが、メモリ46内の音声テンプレートのライブラリと比較され、その音声

発声の内容が判別される。その場合、プロセッサ 40 は、他のデータへのアクセス権を制御するか、特定の活動を実現するか、または周辺システムを口頭制御するために使用できる出力信号を生成してもよい。

【0021】上記の説明を参照し、当業者は、呼出し側の ID を判別できないような状況で使用するために標準化した一連のテンプレートを提供するために中核ライブラリ 48 を使用できること、あるいは、特定の地理的位置固有の語彙項目または通信チャンネル内の伝送パラメータによって大幅に影響を受ける発声を反映するために、中核ライブラリ 48 が、呼出し側固有ライブラリと結合された一連の基本語彙テンプレートを含んでもよいことに留意されたい。いずれの場合にも、プロセッサ 40 は、呼出し側 ID 44 を使用して呼出し側固有ライブラリを選択することで、メモリ 46 内のライブラリと照らし合わせて入力発声を処理し、それにより、コンピュータ 16 内に実現された音声認識システムの効率と精度が大幅に向上する。

【0022】音声認識においてこれまでに知られている試みに関して前述した通り、入力音声の処理に使用するテンプレートは、個々の単語、語句、または単語の一部を表すテンプレートを含むことができる。本明細書で使用する「テンプレート」という用語は、未知の音声発声を識別するためにプロセッサ 40 が使用できるデジタル表現を格納したものを意味する。

【0023】最後に図 3 を参照して説明すると、同図には、本発明の方法を実現するためのプロセスを示す高水準論理流れ図が示されている。図示の通り、このプロセスはブロック 60 から始まり、ブロック 62 に移行する。ブロック 62 は、公衆交換電話網を介してホスト位置への接続をユーザが確立することを示している（図 1 参照）。次に、プロセスはブロック 64 に移行する。ブロック 64 は、言葉の発声が検出されたかどうかの判別を示している。検出されなかった場合は、発声が検出されるまでプロセスが反復するだけである。しかし、言葉の発声が検出されると、プロセスはブロック 66 に移行する。

【0024】ブロック 66 は、呼出し側の ID が電話網から判別できるかどうかの判別を示している。当業者は、呼出し側の ID は普遍的に適用できるものではなく、そのため、システムにアクセスする特定の呼出し側の ID が判別できない場合もあることに留意されたい。ただし、呼出し側の ID が判別された場合、プロセスはブロック 68 に移行する。ブロック 68 は、メモリ 46 内から特定の呼出し側固有ライブラリを選択することを示している（図 2 参照）。前述の通り、特定の呼出し側固有ライブラリは、通信チャンネル内の口頭通信に影響する伝送パラメータを正確に反映するよう処理された音声発声を含むことが好ましい。このため、帯域制限、処理技術、および口頭通信に影響するその他のパラメータを

使用して、そのチャンネルで処理された発声をもっと正確に反映するような音声テンプレートが作成されている。しかも、その呼出し側の ID に関連する特定の地理的地域に特有の具体的な語彙単語と発音も、呼出し側固有ライブラリ内に含まれる。たとえば、定期航空路線の運航スケジュールに口頭アクセスできるようなシステムは、電話網を使用して判別される呼出し側の位置の付近にある地理的位置の発声表現を、その位置からかなり離れたところにある地理的位置より可能性の高い認識用選択項目として含むように設計された、一連の呼出し側固有テンプレートを有することが好ましい。

【0025】呼出し側 ID の判別に基づいて特定の呼出し側固有ライブラリを選択した後で、プロセスはブロック 70 に移行する。ブロック 70 は、その呼出し側固有ライブラリと照らし合わせて発声を処理することを示している。当然のことながら、前述の通り、呼出し側固有テンプレート・ライブラリと照らし合わせて入力音声発声を処理する場合、公衆交換電話網内の通信チャンネルによる伝送の結果として発生する劣化にもかかわらず、容易に認識できる共通発声からなる中核ライブラリと照らし合わせてその発声を処理することを含んでもよい。

【0026】もう一度ブロック 66 を参照すると、呼出し側の ID を判別できない場合は、プロセスはブロック 72 に移行する。ブロック 72 は、呼出し側の ID を判別できないような状況に使用できる中核ライブラリと照らし合わせて入力音声発声を処理することを示している。その後、中核ライブラリまたは呼出し側固有ライブラリを単独でまたは何らかの基本中核ライブラリとともに使用して突合せを行って入力音声発声を処理した後で、プロセスはブロック 74 に移行する。ブロック 74 は、発声が認識されたかどうかの判別を示している。発声が認識されない場合、すなわち、発声と既知のテンプレートとの間に確率の高い一致が存在しない場合、プロセスはブロック 76 に移行する。ブロック 76 は、適当なエラー・メッセージの生成を示しており、その後、プロセスはブロック 78 に移行して復帰する。当業者は、この時点で通信を終了するか、または最終的にその音声の認識に到達できるようにもう一度発声の発音を試みよう話し手に催促できることに留意されたい。

【0027】もう一度ブロック 74 を参照すると、発声が認識された場合は、プロセスはブロック 80 に移行する。ブロック 80 は、その発声の処理を示している。当業者は、「発声の処理」が、その他のデータにアクセスするか、何らかの機能を実行するか、または何らかの方法でコンピュータ 16 の周辺システムと対話して、発声に対して高機能反応を提供するか、それを文字に置き換えるために、その発声の言語内容または情報内容を使用することを意味することに留意されたい。その後、プロセスはブロック 78 に移行して復帰する。

【0028】上記の説明を参照し、当業者は、複数の呼

出し側固有音声テンプレート・ライブラリ内の音声認識システムと、既存の呼出し側識別システムとを組み合わせることで、電話網内の通信チャネルの結果として発生する信号変化を考慮した処理と、地理に固有の語彙および発音およびその他の言語現象の変化を考慮した処理によって入力音声をもっと正確に認識できるため、効率と精度が大幅に向上した音声認識システムを提供できることに留意されたい。

【0029】まとめとして、本発明の構成に関して以下の事項を開示する。

【0030】(1) 電話網を介して音声を表す格納済みテンプレートと照らし合わせて音声の入力フレームを処理する音声認識システムの精度と効率を高めるための方法において、音声テンプレートの中核ライブラリを作成して格納するステップと、複数の呼出し側固有音声テンプレート・ライブラリを作成して格納するステップと、前記電話網内の呼出し側のIDを判別しようと試みるステップと、前記電話網内の前記呼出し側のIDが判別されない場合に、音声テンプレートの前記中核ライブラリと照らし合わせて入力音声発声処理するステップと、前記電話網内の前記呼出し側のIDの判別に対する応答として、前記複数の呼出し側固有音声テンプレート・ライブラリから選択されたライブラリと照らし合わせて入力音声発声処理するステップとを含む方法。

(2) 前記電話網内の呼出し側のIDを判別する前記ステップが、前記電話網内の呼出し側識別システムを使用して前記呼出し側のIDを判別するステップを含むことを特徴とする、上記(1)に記載の電話網を介して音声を表す格納済みテンプレートと照らし合わせて音声の入力フレームを処理する音声認識システムの精度と効率を高めるための方法。

(3) 複数の呼出し側固有音声テンプレート・ライブラリを作成して格納する前記ステップが、前記電話網内での伝送の結果として発生する音声発声の変化を反映するように処理される複数の呼出し側固有音声テンプレート・ライブラリを作成して格納するステップを含むことを特徴とする、上記(1)に記載の電話網を介して音声を表す格納済みテンプレートと照らし合わせて音声の入力フレームを処理する音声認識システムの精度と効率を高めるための方法。

(4) 複数の呼出し側固有音声テンプレート・ライブラリを作成して格納する前記ステップが、具体的な地理的位置を反映する語彙および発音を含む複数の呼出し側固有音声テンプレート・ライブラリを作成して格納するステップを含むことを特徴とする、上記(1)に記載の電話網を介して音声を表す格納済みテンプレートと照らし合わせて音声の入力フレームを処理する音声認識システムの精度と効率を高めるための方法。

(5) 電話網を介して音声を表す格納済みテンプレートと照らし合わせて音声の入力フレームを処理する音声認

識システムの精度と効率を高めるためのシステムにおいて、音声テンプレートの中核ライブラリを作成して格納するための手段と、複数の呼出し側固有音声テンプレート・ライブラリを作成して格納するための手段と、前記電話網内の呼出し側のIDを判別しようと試みるための手段と、前記電話網内の前記呼出し側のIDが判別されない場合に、音声テンプレートの前記中核ライブラリと照らし合わせて入力音声発声処理するための手段と、前記電話網内の前記呼出し側のIDの判別に対する応答として、前記複数の呼出し側固有音声テンプレート・ライブラリから選択されたライブラリと照らし合わせて入力音声発声処理するための手段とを含むシステム。

(6) 前記電話網内の呼出し側のIDを判別するための前記手段が、前記電話網内の呼出し側識別システムを使用して前記呼出し側のIDを判別するための手段を含むことを特徴とする、上記(5)に記載の電話網を介して音声を表す格納済みテンプレートと照らし合わせて音声の入力フレームを処理する音声認識システムの精度と効率を高めるためのシステム。

(7) 複数の呼出し側固有音声テンプレート・ライブラリを作成して格納するための前記手段が、前記電話網内での伝送の結果として発生する音声発声の変化を反映するように処理される複数の呼出し側固有音声テンプレート・ライブラリを作成して格納するための手段を含むことを特徴とする、上記(5)に記載の電話網を介して音声を表す格納済みテンプレートと照らし合わせて音声の入力フレームを処理する音声認識システムの精度と効率を高めるためのシステム。

(8) 複数の呼出し側固有音声テンプレート・ライブラリを作成して格納するための前記手段が、具体的な地理的位置を反映する語彙発音を含む複数の呼出し側固有音声テンプレート・ライブラリを作成して格納するための手段を含むことを特徴とする、上記(5)に記載の電話網を介して音声を表す格納済みテンプレートと照らし合わせて音声の入力フレームを処理する音声認識システムの精度と効率を高めるためのシステム。

【0031】

【発明の効果】本発明によれば、改良された音声認識システムが提供される。また、音声認識の精度と効率を高めるための改良された方法およびシステムが提供される。さらに、呼出し側固有音声テンプレート・ライブラリを使用して電話網内の音声認識を高めるための改良された方法およびシステムが提供される。

【図面の簡単な説明】

【図1】本発明の方法およびシステムを実現するために使用できる分散電話網の絵画図である。

【図2】図1のホスト位置内の音声認識システムの高水準ブロック図である。

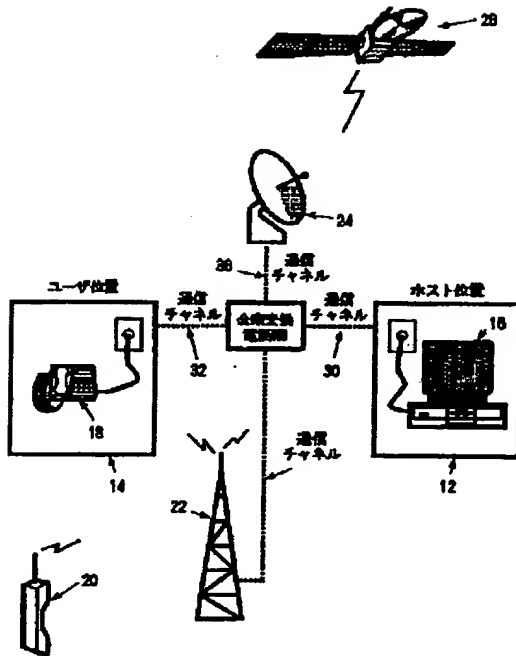
【図3】本発明の方法を実現するためのプロセスを図示する高水準論理流れ図である。

【符号の説明】

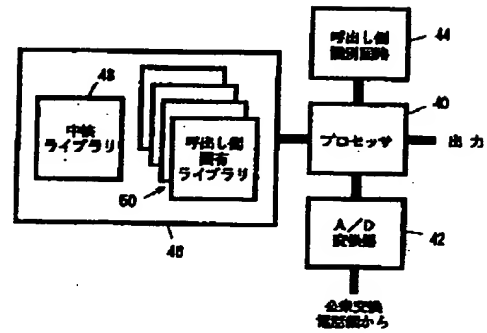
16 コンピュータ
40 プロセッサ
42 A/D変換器

44 呼出し側識別回路
46 メモリ
48 中核ライブラリ
50 呼出し側固有ライブラリ

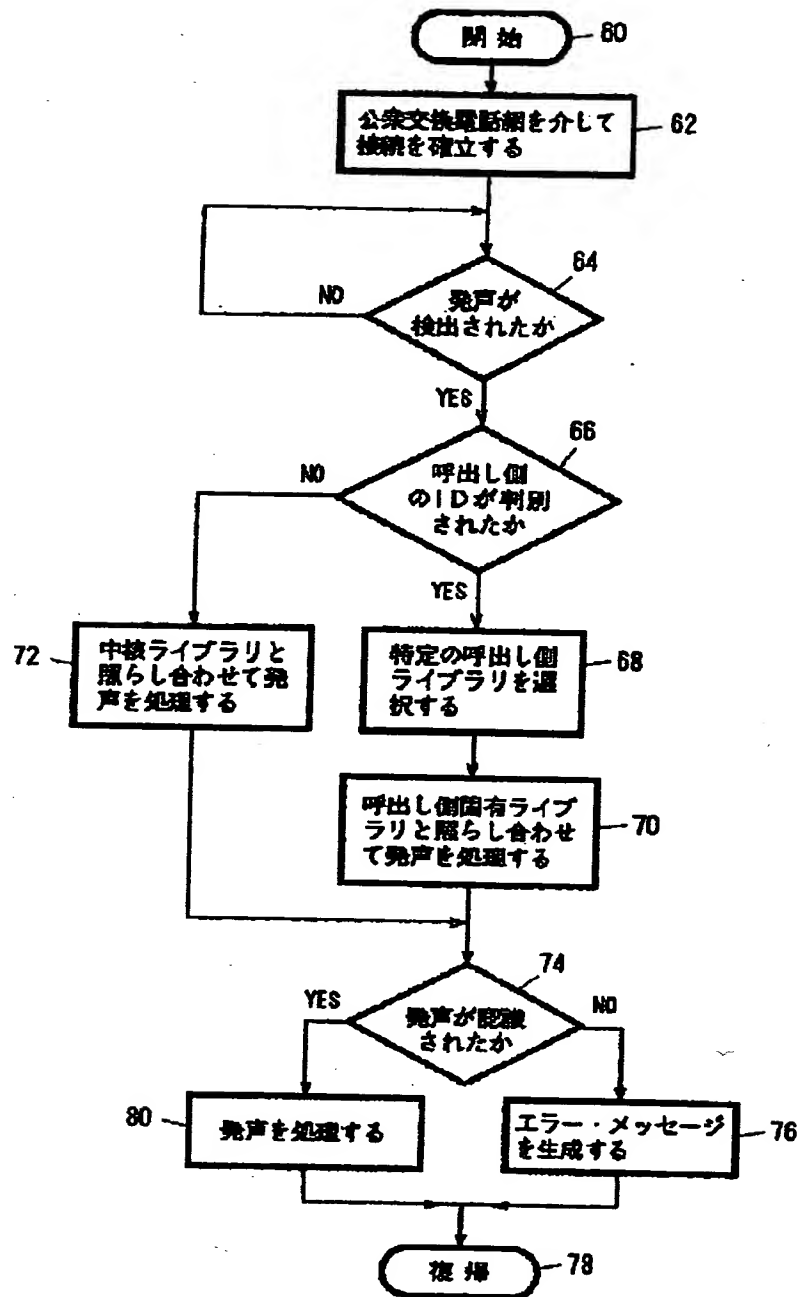
【図1】



【図2】



【図 3】



フロントページの続き

(72)発明者 ジョン・エム・リュカッセン
 アメリカ合衆国10025 ニューヨーク州ニ
 ューヨーク ワンハンドレッドアンドサー
 ド・ストリート ウェスト 308

(72)発明者 エルトン・ビー・シャーウィン・ジュニア
 アメリカ合衆国06903 コネチカット州ス
 タンフォード ドッグウッド・レーン 26

(72)発明者 ホルヘ・エル・ビスカイノ
アメリカ合衆国33067 フロリダ州パーク
ランド セブンティーセカンド・ブレース
ノースウエスト 6442

THIS PAGE BLANK (USPTO)